capturing) and stretches to fill the base of the view frustum. In some implementations, pose correction can include three operations: background detection, extrusion, and rotation.

In some implementations, background detection can include using depth data to detect or determine the background. There are two main technical problems with background detection using live depth frames: noise and bleeding effect (e.g., halo like artifact due to missing data near the edges). Some implementations provide a solution to the bleeding effect problem by applying a Sobel filter to detect edges of areas that can be treated as background. To address the noise problem, such as spatial or temporal noise, a smoothing algorithm can be used.

A shader can be implemented for the above functionality returning a value in a range (e.g., [0,1]) defining the confidence of whether a given point is background or foreground (thus providing information that can be used for a fade effect around the edges of the foreground portion). The shader can include a per-pixel classifier that executes runs on a GPU instead of a CPU, to achieve efficiency. The shader receives, as input, the color image and produces a foreground/background mask as an output image (e.g., red=foreground and black=background, etc.). Another shader performs the pose correction using the color image, depth image, foreground/background mask, and/or camera intrinsics (e.g., focal length of the camera, and other camera parameters or settings) as input.

Some implementations can include extrusion and rotation of at least a portion of one or more video frames (or other images). The extrusion and rotation can include using smoothed disparity map data. The disparity map data can be applied to the foreground portion of the image (in the camera-plane representation) to move foreground pixels closer to the camera. A disparity map is another way to represent depth data (e.g., representing distance of the device camera to a point).

In some implementations, a virtual camera position is updated to be perpendicular to the face plane in one or more frames (e.g., in the first frame). Updating the virtual camera position can include setting the forward vector of the camera to be the opposite of forward vector of the face anchor. A face anchor includes face position and face rotation angle calculated during face detection. In some implementations, the projection matrix of the camera can be updated to go through the four edges of the image plane using a calculation that is done whenever the device moves and that takes into account a current delta between the face pose and a "front pose" in order to not cancel out intentional face movements. In some implementations, the projection matrix can be maintained for foreground pixels and the projection for background pixels can be skipped to reduce computation time, with a caveat being that foreground pixels are visible. The foreground can be zoomed in to hide any imperfect matches with the background such that the foreground occludes the background and is then enlarged to cover any gaps created by the face rotation.

Some implementations can include background treatments. In some implementations, background treatments can include adjusting the background, which could be stretched due to rotation. In order for the background to not have a distracting appearance, a blurring effect can be used on the background.

The systems and methods provided herein may overcome one or more deficiencies of some conventional video calling systems and methods. For example, some conventional video calling system may not provide a capability to detect, correct or adjust for a camera angle that may contribute to capturing a lower quality video or may cause other problems with video frames such as exaggerated features and/or movement effects. With the easy availability of digital image capture devices, such as digital cameras, phones with built-in cameras, wearable devices with cameras, head-mounted devices, tablets, personal computers, etc., users may frequently capture or transmit video.

The example systems and methods described herein may overcome one or more of the deficiencies of conventional video capture systems to provide users with automatic pose correction of videos or other images. A technical problem of some conventional systems may be that such systems do not detect and correct poses based on image analysis and corresponding confidence scores. The disclosed subject matter relates to particular techniques to correct poses of videos or other images and/or to automatically correct for other video or image issues. Some implementations of the disclosed subject matter provide a solution to one or more of the above-mentioned technical problems. For example, by re-projecting video, some implementations enable high quality video irrespective of camera angle.

Videos are used herein as examples to describe the subject matter. Other types of images could be used with an implementation of the disclosed subject matter. An image may be a static image (e.g., a single frame with no motion), an animated image, a video (e.g., with a plurality of frames), etc. For example, a static image may depict one or more objects, while an animated image may depict one or more objects that change within the image (e.g., a live photo that captures a face with eyes transitioning between closed and open, face with the mouth moving from a non-smiling to a smiling position, etc.) A video may include a plurality of frames that depict one or more persons or objects. Pose correction may be provided for any type of image, and may include a group of images of same and/or different types.

FIG. **1** is a diagram showing an example user **102** using a user device **104** to conduct a video call. The user device **104** is at an angle relative to the user's face that is capturing video of the user at an angle that may produce a lower quality video, such as the example lower quality video frame shown in FIG. **3**.

Due to the angle of the user device **104** relative to the face of the user **102**, a first distance **106** from the camera to the top of the user's head is different (e.g., longer) than a second distance **108** from a camera of the user device **104** to the bottom of the user's **102** face. A result of the difference in the distances **106** and **108** resulting from the upward angle of the user device **104** creates a pose (e.g., as shown on the device **302** in FIG. **3**) that may not provide a high quality video (e.g., may cause distortion, etc.).

In addition to the pose at an angle that may produce a lower quality video, other problems can be introduced in the video such a movement effects (e.g., due to hand movement or jitter by the user **102**) and image effects (e.g., the big nose effect) due to the lens of the user device **104**. These problems may result from the physical placement of the camera on a user device (e.g., at the top of the device) and from a common posture for holding user devices such as mobile phones or wearable devices, e.g., smartwatch or other device, in which the hand holding the phone (or other device) is at a downward angle relative to the face of the user. In addition to distortion of the foreground portions of the images (e.g., a pose captured at an angle that may produce a lower quality video **304**), the camera angle can also result in a background **306** that is distorted.

FIG. **2** is a diagram showing the example user device **104** being used to conduct a video call with the user device **104**